We claim:

1. A computer-implemented method for constructing a single vector representing a semantic abstract in a topological vector space for a semantic content of a document on a computer system, the method comprising:

5   storing a semantic content for the document in computer memory accessible by the computer system;

constructing state vectors in the topological vector space for the semantic content;

superpositioning the state vectors to construct the single vector; and

storing the single vector as the semantic abstract for the document.

10

2. A method according to claim 1, wherein constructing the state vectors includes:

identifying lexemes/lexeme phrases in the semantic content; and

constructing a state vector for each lexeme/lexeme phrase in the semantic content.

15

3. A method according to claim 1, wherein superpositioning the state vectors includes adding the state vectors using vector arithmetic.

4. A method according to claim 1, wherein superpositioning the state vectors

20   includes weighting the state vectors.

5. A method according to claim 1 further comprising normalizing the single vector.

25   6. A method according to claim 1, wherein:

storing a semantic content includes:

storing the document in computer memory accessible by the computer system; and

extracting words from at least a portion of the document;

30   constructing state vectors includes constructing a state vector in the topological vector space for each word using a dictionary and a basis; and
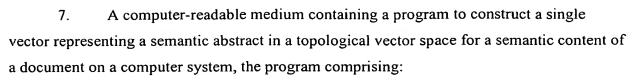
the method further comprises filtering the state vectors.

7.   A computer-readable medium containing a program to construct a single vector representing a semantic abstract in a topological vector space for a semantic content of a document on a computer system, the program comprising:

storing software to store a semantic content for the document in computer memory accessible by the computer system;

construction software to construct state vectors in the topological vector space for the semantic content;

superpositioning software to superposition the state vectors to construct the single vector; and

storing software to store the single vector as the semantic abstract for the document.

8.   A program according to claim 7, wherein the construction software includes:

identification software to identify lexemes/lexeme phrases in the semantic content; and

construction software to construct a state vector for each lexeme/lexeme phrase in the semantic content.

9.   A program according to claim 7, wherein the superpositioning software includes addition software to add the state vectors using vector arithmetic.

10.   A program according to claim 7, wherein the superpositioning software includes weighting software to weigh the state vectors.

11.   A program according to claim 7 further comprising normalization software to normalize the single vector.

12.   A program according to claim 7, wherein:

the storing software includes:

storing software to store the document in computer memory accessible by the computer system; and

extraction software to extract words from at least a portion of the document;

the construction software includes construction software to construct a state vector in the topological vector space for each word using a dictionary and a basis; and

the program further comprises filtering software to filter the state vectors.

13. An apparatus on a computer system to construct a single vector representing a semantic abstract in a topological vector space for a semantic content of a document on a computer system, the apparatus comprising:

a semantic content stored in a memory of the computer system;

a state vector constructor for constructing state vectors in the topological vector space for the semantic content; and

a superpositioning unit adapted to superposition the state vectors into a single vector as the semantic abstract.

14. An apparatus according to claim 13, wherein:

the state vector includes an associated threshold distance; and

the apparatus further comprises:

search means for searching the topological vector space for a second document with a second semantic abstract within the threshold distance associated with the first semantic abstract for the first document; and

retrieval means to retrieve the second document.

15. An apparatus according to claim 13, wherein the state vector constructor includes a lexeme identifier adapted to identify lexemes/lexeme phrases in the semantic content.

16. An apparatus according to claim 13, wherein the superpositioning unit includes a vector arithmetic unit adapted to add the state vectors.

17. An apparatus according to claim 13 further comprising a normalization unit adapted to normalize the single vector.

18. An apparatus according to claim 13, wherein:

the apparatus further comprises:

a lexeme extractor adapted to extract lexemes/lexeme phrases from the semantic content; and

MJM Docket No. 6647-16

filtering means for filtering the state vectors; and

the state vector constructor is adapted to constructing a state vector in the topological vector space for each lexeme/lexeme phrase using a dictionary and a basis.

5   19.   A computer-implemented method for constructing minimal vectors representing a semantic abstract in a topological vector space for a semantic content of a document on a computer system, the method comprising:

storing a semantic content for the document in computer memory accessible by the computer system;

10   constructing state vectors in the topological vector space for the semantic content;

locating clumps of state vectors in the topological vector space;

superpositioning the state vectors within each clump to form a single vector representing the clump;

collecting the single vectors representing each clump to form the minimal vectors; and

15   storing the minimal vectors as the semantic abstract for the document.

20.   A computer-readable medium containing a program to construct minimal vectors representing a semantic abstract in a topological vector space for a semantic content of a document on a computer system, the program comprising:

20   storing software to store a semantic content for the document in computer memory accessible by the computer system;

construction software to construct state vectors in the topological vector space for the semantic content;

clump location software to locate clumps of state vectors in the topological vector

25   space;

superpositioning software to superposition the state vectors within each clump to form a single vector representing the clump;

collection software to collect the single vectors representing each clump to form the minimal vectors; and

30   storing software to store the minimal vectors as the semantic abstract for the document.

MJM Docket No. 6647-16

21.   An apparatus on a computer system to construct minimal vectors representing a semantic abstract in a topological vector space for a semantic content of a document on a computer system, the apparatus comprising:

a semantic content stored in a memory of the computer system;

a state vector constructor for constructing state vectors in the topological vector space for the semantic content;

a clump locator unit adapted to locate clumps of state vectors in the topological vector space;

a superpositioning unit adapted to superposition the state vectors within each clump into a single vector representing the clump; and

a collection unit adapted to collect the single vectors representing the clump into the minimal vectors of the semantic abstract.